What is claimed is:

1.   A method for receiving a video sequence including query
objects to be extracted and generating object-labeled images based on the
query objects, the method comprising the steps of:

(a) dividing the video sequence into one or more shots, each of which
is a set of frames having a similar scene, and selecting one or more key
frames from each of the shots;

(b) extracting query object based initial object regions from each of
the key frames;

(c) tracking object regions in all frames of each of the shots based on
the corresponding query image based initial object regions; and

(d) labeling the object regions tracked in each of the frames based on
information on the corresponding query objects.

2.   The method of claim 1, wherein step (b) comprises:
(b1) determining whether there exists an object similar to each of the query
objects in each of the key frames, and if there is a similar object in a key
frame, extracting the similar object as a corresponding query object based
initial object region; and

(b2) generating query object based shot mask images in all key
frames of the shots by setting pixels of the query object based initial object
regions extracted from each of the key frames as a first value and setting
the remaining pixels of each of the key frames as a second value.

3.   The method of claim 2, wherein step (c) comprises:
(c1) tracking the object regions in all frames of each of the shots based on
the corresponding query image based shot mask images and video feature
values of the corresponding query objects; and

(c2) generating query object based frame mask images in all frames

13

6   of each of the shots by setting pixels of the object regions tracked in each of

7   the frames as a first value and setting the remaining pixels of each of the

8   key frames as a second value.

1   4.   The method of claim 3, wherein, in step (d), each of the object

2   regions is labeled in each frame with a unique number set to the

3   corresponding query image or coordinate information of the corresponding

4   query image in each frame.

1   5.   An apparatus for receiving a video sequence including query

2   objects to be extracted and generating object-labeled images based on the

3   query objects, the apparatus comprising:

4   a shot and key frame setting unit for dividing the video sequence into

5   one or more shots, each of which is a set of frames having a similar scene,

6   and selecting one or more key frames from each of the shots;

7   an initial object region extractor for extracting query object based

8   initial object regions from each of the key frames;

9   an object region tracker for tracking object regions in all frames of

10   each of the shots based on the corresponding query image based initial

11   object regions; and

12   an object-labeled image generator for labeling the object regions

13   tracked in each of the frames based on information on the corresponding

14   query objects.

1   6.   The apparatus of claim 5, wherein the initial object region

2   extractor determines whether there exists an object similar to each of the

3   query images in each of the key frames, and if there is a similar object in a

4   key frame, extracts the similar object as a corresponding query object based

5   initial object region, and generates query object based shot mask images in

14

6   all key frames of each of the shots by setting pixels of the query object

7   based initial object regions extracted from each of the key frames as a first

8   value and setting the remaining pixels of each of the key frames as a

9   second value.

1   7.   The apparatus of claim 6, wherein the object region tracker

2   tracks the object regions in all frames of each of the shots based on the

3   corresponding query image based shot mask images and video feature

4   values of the corresponding query objects, and generates query object

5   based frame mask images in all frames of each of the shots by setting pixels

6   of the object regions tracked in each of the frames as a first value and

7   setting the remaining pixels of each of the key frames as a second value.

1   8.   The apparatus of claim 5, wherein the object-labeled image

2   generator labels each of the object regions in each frame with a unique

3   number set to the corresponding query image or coordinate information of

4   the corresponding query image in each frame.

1   9.   A computer readable medium having embodied thereon a

2   computer program for receiving a video sequence including query objects to

3   be extracted and generating object-labeled images based on the query

4   objects, wherein generating object-labeled images comprises the steps of:

5   (a) dividing the video sequence into one or more shots, each of which

6   is a set of frames having a similar scene, and selecting one or more key

7   frames from each of the shots;

8   (b) extracting query object based initial object regions from each of

9   the key frames;

10   (c) tracking object regions in all frames of each of the shots based on

11 the corresponding query image based initial object regions; and

15

12   (d) labeling the object regions tracked in each of the frames based on

13   information on the corresponding query objects.